

Having thus described our invention, what we claim as new and desire to secure by letters patent is:

1. An apparatus performing parallel binary-binary matrix-vector multiplication with embedded storage of the matrix; the apparatus comprising an array of charge-based cells receiving binary inputs, storing binary matrix elements and returning analog outputs; each cell comprising:

A first device storing charge representing one said binary matrix element, the stored charge coupling capacitively to an output line;

A second device coupled to said first device, where transfer of said charge between said first and second device in a computation cycle is controlled by an input line;

A third device coupled to said first device and to a data line, where write or refresh of said charge is activated onto said data line through a select line.

2. The apparatus recited in claim 1 wherein said first, second and third device in said charge-based cell comprise field effect transistors.

3. The apparatus recited in claim 1 further comprising circuits assisting in write and dynamic refresh of said charge in said charge-based cells.

4. The apparatus recited in claim 1 wherein said analog outputs are converted to digital outputs through quantization.

5. The apparatus recited in claim 1 performing digital-digital matrix-vector multiplication; the apparatus comprising said array of charge-based cells receiving bit-serial digital inputs over multiple computation cycles, storing bit-parallel matrix elements spanning multiple rows of said array, and returning analog or digital outputs combining analog or quantized outputs from said array over said computation cycles and said rows.

6. The apparatus recited in claim 1 performing parallel signed binary-binary matrix-vector multiplication with embedded storage of the matrix; the apparatus comprising an array of complementary cells receiving complementary signed binary inputs, storing complementary signed binary matrix elements and returning analog outputs; each complementary cell comprising two said charge-based cells; each charge-based cell receiving one polarity of said input and storing one polarity of said matrix element.

7. The apparatus recited in claim 6 wherein said analog outputs are converted to digital outputs through quantization.

8. The apparatus recited in claim 6 performing signed digital-digital matrix-vector

multiplication; the apparatus comprising said array of complementary cells receiving complementary bit-serial digital inputs over multiple computation cycles, storing complementary bit-parallel matrix elements spanning multiple rows of said array, and returning analog or digital outputs combining analog or quantized outputs from said array over said computation cycles and said rows.

9. A method for large-scale high-resolution digital matrix-vector multiplication using a parallel signed binary-binary matrix-vector multiplier; said matrix-vector multiplier receiving signed binary inputs, storing signed binary matrix elements and returning analog outputs; the method comprising:

modulation of digital inputs to produce pseudo-random inputs;

signed bit-serial presentation of said pseudo-random inputs to said signed binary-binary matrix-vector multiplier;

quantization of corresponding analog outputs to produce partial digital outputs;

combination of said partial digital outputs to produce pseudo-random digital outputs;

demodulation of said pseudo-random digital outputs to undo the effect of said modulation of said digital inputs, producing desired digital outputs.

10. The method of claim 9 using a parallel signed digital-binary matrix-vector multiplier; said matrix-vector multiplier receiving signed binary inputs, storing digital matrix elements in signed bit-parallel form over multiple rows, and returning analog outputs; said combination of said partial digital outputs spanning said multiple rows.

11. The method of claim 10 wherein said digital inputs are modulated by digitally subtracting reference inputs drawn from a random distribution to produce said pseudo-random inputs, and wherein said pseudo-random digital outputs are demodulated by digitally adding the result of multiplying said digital matrix with said reference inputs to produce said desired digital outputs.

12. The method of claim 11 wherein said result of multiplying said digital matrix with said reference inputs is obtained from said digital-binary matrix multiplier.

13. The method of claim 11 wherein said reference inputs are fixed, and wherein said result of multiplying said digital matrix with said reference inputs is precomputed and stored.